## Chapter 6 "Dealing With Data" SPSS Tutorial

1. **Visit** the student website at for this textbook at www.clowjames.net/students.

2. **Download** the following files:
   "Chapter 06 Dealing with Data" (SPSS data file)
   "Chapter 06 Dealing with Data Variables" (Word file)

3. Open SPSS 17.0 then open the data file named, *"Chapter 06 Dealing with Data"*. (If you need help, review the Dealing With Data instructions from chapter 2.)

4. Read the "Statistics Review" section of chapter 6 if you haven't already done so.

5. In Microsoft Word, open the "Chapter 06 Dealing with Data Survey" to review the survey questions.

One-Sample T-Tests

6. From the top menu bar in SPSS, select ANALYZE > COMPARE MEANS > ONE SAMPLE T TEST.
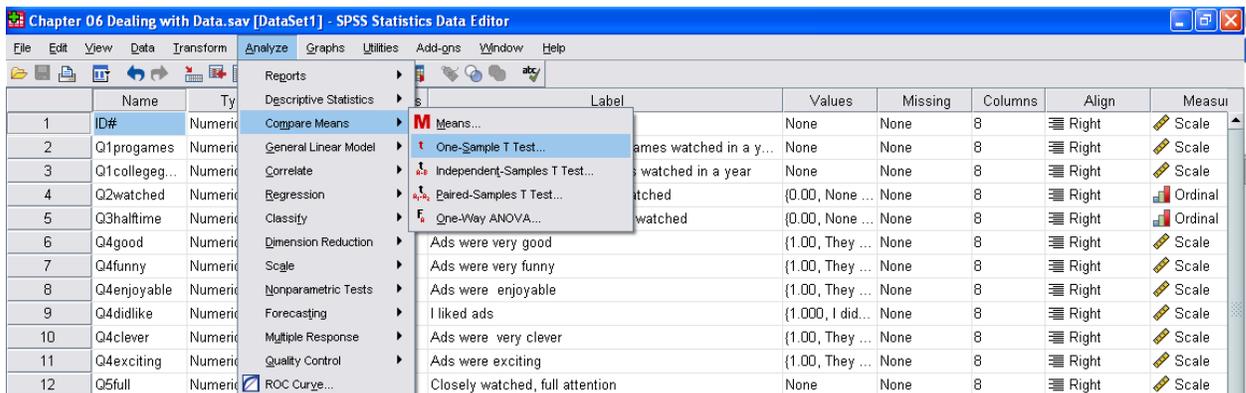

   The pop-up box shown below will appear.

7. From the left-hand box of variables, highlight the variable to be tested. You will choose the third variable, which lists the **number of college football games watched in a year** for each respondent.

8. Once you have highlighted the appropriate variable, click the arrow button to move the variable to the TEST VARIABLE(S) box on the right hand side of the One-Sample T Test pop-up box.
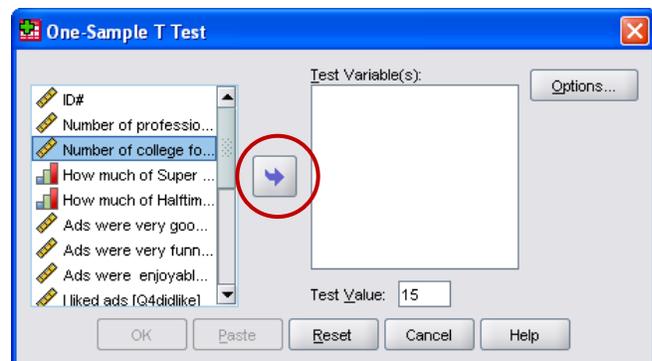
9. Next, enter the TEST VALUE in the appropriate box.   Recall that the null hypothesis being tested is that the average number of college football games watched annually is equal to 15. This represents the **minimum** number of college football games that **must be watched** on average in order for the businesswoman to invest in advertising.
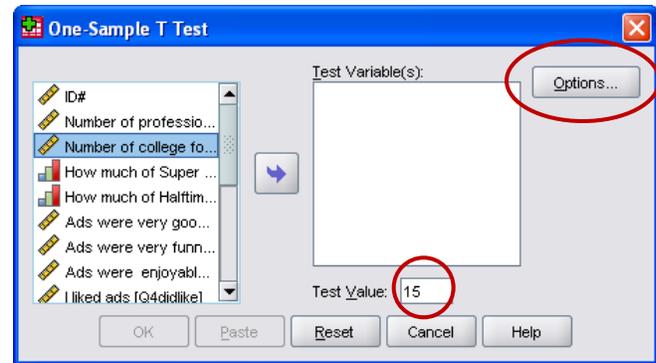
10. Now click the OPTIONS button. The "One-Sample T Test: Options" dialogue box will appear. The confidence interval defaults to 95%. However, you also need to run tests at the 90% and 99% confidence interval level. Change this number to 90 (%). You will need to repeat the t-test twice after your first run to test the 95% and 99% confidence intervals. You cannot test different confidence levels at the same time.
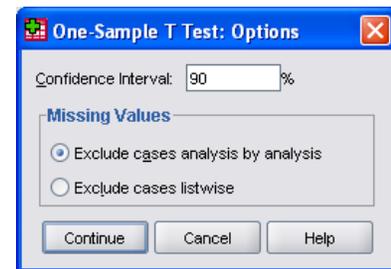
11. As the exercise asks you to test only a single variable, you can leave the MISSING VALUES option set to "exclude cases analysis by analysis."

12. Once you have entered the confidence interval and the missing values information, click the CONTINUE button, double check to be certain that you have entered the correct test variable and test value on the remaining dialogue box, then click OK.

13. Repeat this process twice more, making certain that you perform the one sample t-test at the 90%, 95%, and 99% confidence level. Most of the data will stay the same, with the exception of the lower and upper numbers found under the confidence interval of the difference portion of table 2. Review the Statistics Tutorial on pages 183-185 for help with interpreting the one sample t-test results.

Analysis of Variance (ANOVA)

14. From the top menu bar in SPSS, select ANALYZE > COMPARE MEANS > ONE WAY ANOVA.
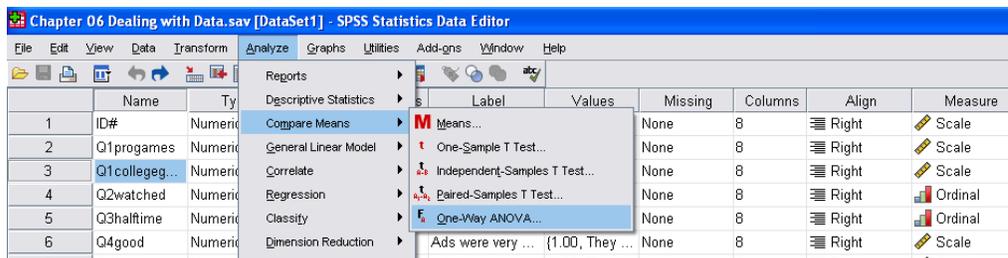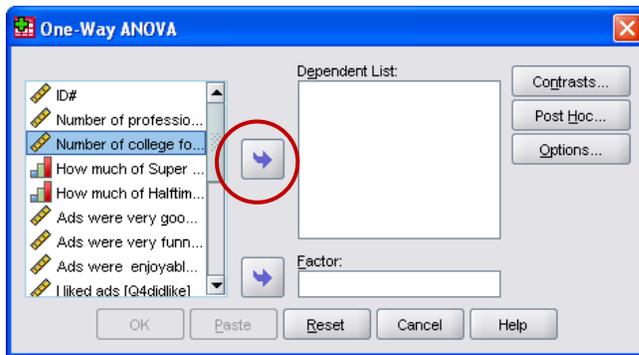
15. The One-Way ANOVA pop-up box will appear as shown at left.

16. From the left-hand box of variables, highlight the DEPENDENT VARIABLE to be tested. The dependent variable is the one for which mean scores will be calculated. You will choose the third variable, which lists the **number of college football games watched in a year** for each respondent.

17. Once you have highlighted the appropriate variable, click the arrow button to move the variable to the DEPENDENT LIST box on the right hand side of the One-Way ANOVA pop-up box.

18. Next, select the FACTOR variable. The Factor variable is also known as the grouping variable. Grouping or factor variables are typically demographic or categorical in nature. Gender, age, classification, and race could all be used as factor variables in an ANOVA, as could questions 2 through 4.
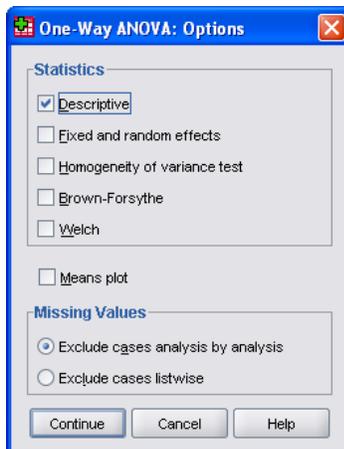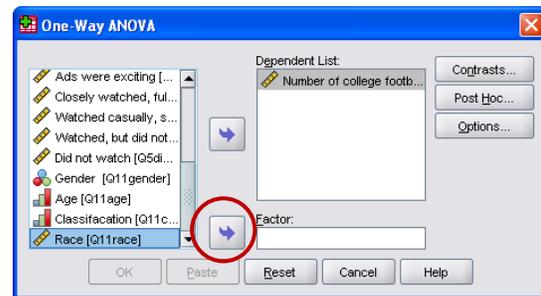
19. Click the OPTIONS button which is located to the right of the dependent list. Selecting DESCRIPTIVES is helpful, as the resulting table will show the mean, standard deviation, standard error and sample size for each group tested by the ANOVA.

20. As you examine the DESCRIPTIVES out, you'll notice that the sample mean values for the number of college games watched by each race as displayed in the Descriptives table are different from one another.

21. However, **you must use the results of the ANOVA test** to determine whether or not the differences in the number of college games found in the sample data is likely to represent real differences in the **population**. The null hypothesis being tested is that there is no difference between the number of college games watched by each group. IF the p-value shown under the Sig. column is less than or equal to the test value (.05), the null hypothesis can be rejected as the probability is very high that the differences found in the sample are also likely to be found in the population.

ANOVA Post-Hoc Tests

The following steps should be followed ONLY if the overall ANOVA test rests were significant for factor variables containing three or more groups.

22. From the top menu bar in SPSS, select ANALYZE > COMPARE MEANS > ONE WAY ANOVA.

23. If the dependent variable and factor from the previous test are not shown in their respective windows, select the number of college football games as the dependent variable, and race as the factor (review steps 15-18 if you need help).

24. Now click the Post Hoc button which is located to the right of the dependent list on the one-way ANOVA dialogue box. The One-Way ANOVA: Post Hoc Multiple Comparisons dialogue box will appear.
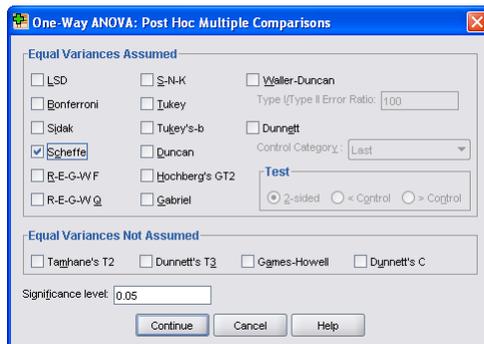
25. Post Hoc Multiple Comparisons allows the researcher to compare the means of each group with every other group to determine which are statistically different from one another.

    This is necessary because a significant ANOVA test result only tells us that differences exist somewhere in the data, but does not tell us which group means are significantly different from one another when three or mean groups are involved.

26. Multiple types of post-hoc comparison tests are available. The most commonly used are the Scheffe, Tukey and Tukey's-b, and Duncan tests. The Scheffe test is often preferred for pairwise comparisons because it is the most conservative test, requiring the largest differences between means for significant results to be found. Thus the chance of making a mistake (finding significant differences in the sample data when it is NOT true for the population) is reduced.

27. Click the box next to the "Scheffe" test.

28. Click the CONTINUE button. Then click OK. Review the output from the test.

Interpreting ANOVA Post-Hoc Tests

29. The output of the Scheffe test will contain descriptive statistics, the ANOVA test table, the post hoc tests and the homogeneous subsets output. Only the last two tables will be discussed here.

Multiple Comparisons

Number of college football games watched in a year
Scheffe

(I) Age	(J) Age	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
18-22	23-29	8.17516*	2.45996	.004	2.1377	14.2126
	30+	5.18343	2.94578	.214	-2.0463	12.4132
23-29	18-22	-8.17516*	2.45996	.004	-14.2126	-2.1377
	30+	-2.99173	3.40093	.679	-11.3385	5.3551
30+	18-22	-5.18343	2.94578	.214	-12.4132	2.0463
	23-29	2.99173	3.40093	.679	-5.3551	11.3385

*. The mean difference is significant at the 0.05 level.

30. An example of post-hoc test results in which AGE is the factor variable is shown above. This table lists the mean difference between each of the groups being compared, the standard error, and the significance level of the post –hoc test.   The null hypothesis being tested is that the mean of group 1 is equal to the mean of group two. This data shows that there is a significant difference in the number of college games watched between those who are 18-22 years old, and those who are 23-29 years old (sig. of .004 < .05). None of the other post-hoc comparison tests were significant.

Number of college football games watched in a year

Scheffe[a,b]

Age	N	Subset for alpha = 0.05 — 1	Subset for alpha = 0.05 — 2
23-29	127	9.8583	
30+	80	12.8500	12.8500
18-22	359		18.0334
Sig.		.600	.217

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 129.536.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

31. The homogenous subset table on the previous page displays the sample size and means for each age group, and visually displays the results of the significance test. Means that only appear in the first column are significantly different from means that only appear in the second column. If two means appear in the same column, we can't be 95% confident that these means are truly different from one another in the population. In more formal terms, we fail to reject the null hypothesis that the means of each group are equal.

32. The age data from our sample suggests that those who are 18-22 years old watch significantly MORE football games than do individuals between the ages of 23 and 29. Specifically, those 18-22 watch on average 18 games a year (column 2), while those 23-29 watch 9.85 games per year (column 1). Those 30 years of age or older watched 12.85 games on average. Remember, this is data from a SAMPLE of individuals, not the population as a whole. So although the mean is numerically different from each of the other groups, the difference in the means between groups is simply NOT large to be 95% confident that it represents real differences in the population as whole.